DOCUMENT RESUME

ED 387 096 IR 017 313

AUTHOR Gibbs, William J.
TITLE Multimedia and Computer-Based Instructional Software: Evaluation Methods.
PUB DATE 95
NOTE 9p.; In: Association of Small Computer Users in Education (ASCUE) Summer Conference. Proceedings (28th, North Myrtle Beach, South Carolina, June 18-22, 1995); see IR 017 305.
PUB TYPE Guides - Non-Classroom Use (055) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; *Computer Assisted Instruction; Computer Software Development; *Courseware; *Evaluation Methods; Hypermedia; Material Development; Multimedia Instruction; *Multimedia Materials

ABSTRACT

Testing and evaluation are important components of effective instructional software development. Many evaluation approaches today are multifaceted, employing a combination of methodologies and multiple data collection techniques. In this paper, several approaches to evaluation are discussed. Computerized tracking systems, videotape recording techniques, and verbal protocol analysis are presented as means by which to conduct more direct comprehensive evaluations of multimedia and hypermedia learning environments, as well as for observing learners' information processing. Results from inquiries that have employed these approaches are discussed, and their potential to affect the study of learners" cognitive processes is reviewed. Information is also provided on cognitive information processing and knowledge construction. Three tables illustrate concepts. (Contains 19 references.) (MAS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

1995 ASCUE Proceedings

Multimedia and Computer-Based Instructional Software: Evaluation Methods

William J. Gibbs
Department of Media Services
Eastern Illinois University
Charleston, IL 61920
217-581-6011
cfwjg1@eiu.edu

ED 387 096

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Introduction

Testing and evaluation are important components of effective instructional software development. While in development, computer-based instructional systems are typically put through a variety of evaluation procedures. There are a host of methods available to researchers to evaluate and study such things as software execution, operating efficiency and the human-computer interface. Many approaches today are multifaceted, employing a combination of methodologies and multiple data collection techniques. These provide researchers greater capability to evaluate and study computer-based instructional systems. Moreover, given developments in cognitive learning theory, they may also be useful in studying learners' cognitive processing while engaged in technology-based instruction.

In this paper, several approaches to evaluation are discussed. Computerized tracking systems, videotape recording techniques, and verbal protocol analysis are presented as means by which to conduct more direct comprehensive evaluations of multimedia and hypermedia learning environments. Results from inquiries that have employed these approaches are discussed and their potential to affect the study of learners' cognitive processes reviewed.

Background

Multimedia is the integration of media (e.g., text, audio, graphics, and video) into a computer-based system (Jonassen, 1995). "It dynamically links and manages organized nodes of information containing multiple symbol systems and images within a given medium and across different media" (Park & Hannafin, 1993, p. 63). Multimedia is often structured as hypermedia (Jonassen, 1995). Hypermedia, characterized by an arrangements of nodes (concepts in text or graphical form) and links (semantic relationships between concepts), allows users non-sequential access to mediated content based on personal need and interest (Kumar, Helgeson & White, 1994). These complex interactive learning systems pose significant challenges for those researching their effects on learning and learners (Gay & Mazur, 1993). For example, the flexibility of a hypermedia system allows for the creation of unique lesson structures reflective of individual learning requirements (Park & Hannafin, 1993) and provides multiple knowledge representations in a variety of media forms (e.g., graphics, sound, animation, and video). While powerful from an information processing perspective, flexible, non-sequential, user-centered designs make it more complex to examine how learners interact with a system (Gay & Mazur, 1993). As Wadlow points out,

> It is difficult to make definitive statements about user-interface design, even in areas which have been studied in great detail, such as text-based processing and command-oriented systems. When this territory is broadened to include multiple windows, color, multiple input devices, and new object types such as animation and video, the user interface design task becomes quite challenging (1990, p. 181).

IR017313

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C.P. Singer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Learners' interactions with a computer-based system is vitally important for design and development. The success of a system is often contingent on an appropriate human-computer interface.

Formative Evaluation

Formative evaluations are iterative in nature whereby program features are evaluated and modified and then reevaluations are subsequently made (Hannafin & Peck, 1988). Methods have been devised to assess the human-computer interface and to collect data about how learners interact with learning systems. Initial program designs, for example, are sometimes given to "experts" for review. Interviews, observations, and pilot testing are other forms of evaluation. Often in pilot testing, students write down their reactions to programs and their feedback is used to refine the courseware. According to Allessi and Trollip (1985), "pilot testing is a process whereby representatives from a target audience use and test lessons while their progress and performance are monitored" (p.384). Various aspects of the systems are evaluated such as program execution, program efficiency, and instructional soundness.

There are several techniques by which to collect data on the human-computer interaction, as well as to study users' cognitive processes. Formative evaluation practices often employ computers to record key stroke data and learner responses: data helpful for evaluating a program and its usefulness to learners. However, in attempts to develop a practical understanding of how computer-based systems are used, recent evaluation approaches have taken a more holistic view (Winograd & Flores, 1986) which advocate qualitative methodologies (Card, Moran & Newell, 1983) using multiple data collection instruments (Marchionini, 1990). Techniques for monitoring the interactions between learner and computer now compile visual, textual, and auditory information. Among other things, this data helps to identify learners' use patterns, reactions, needs, and interactions and portrays, to some extent, a more fuller representation of them in the learning environment. From it, researchers can potentially design more powerful learning environments. Furthermore, in recent years there has been much interest in methods that help researchers understand cognition and the interplay between learners and computer-based learning environments. Since interactive systems designed as hypermedia can, to a degree, resemble learners' knowledge structures, they can be used as devices to gain understanding of processes of cognition (Kumar, Helgeson & White, 1994). A learner's interactions with a computer can be collected using multiple data collection instruments and analyzed to obtain a fuller awareness of their thought processes.

Formative Evaluation Approaches

The discussion to follow will present several potential methods for evaluating learning systems and observing learners' information processing.

Use Pattern Tracking and Student Commentaries

Gibbs and Armel (1994) developed an interactive self-study computer-based module called Imposition designed for students to supplement class lectures. The program provided a visual experience to convey concepts including graphic demonstrations, testing exercises, electronic note taking, on-line glossary, and hypermedia access to key concepts. Also included in the program were two modes of data recording; 1) an on-line tracking system which kept logs of how students navigated through the program; 2) and a facility which enabled students to record comments and notes. All data were written to a file at the end of each program session. Table 1 depicts selected data collected by Imposition's tracking system. The system recorded, 1) the events or concepts a student examined and the sequence in which there were examined (column 1); 2) user number (column 2); 3) number of times events

were examined (column 3); 4) amount of time spent on each event (column 4); 5) students' typed notes (column 5); 6) students' typed comments (column 6) and 7) options (Notes, On-line Glossary or Comments) selected.

Several observations can be made from this data. For example, the student comments in column 6 of Table 1, may indicate that program revisions are necessary to alleviate confusion as to the system's directions on an exercise segment. This is evident by the following student comments, "Are we to number pages? I missed the directions." The designer must determine the reasons for the student missing the directions. Such information is important to alleviate further problems and misunderstandings. The data also reveal that this particular student took a cursory look at the module and spent slightly more time on events in the beginning of the program than those toward the end. Moreover, while an on-line glossary and note taking facility were available, they were not used and only two comments were recorded. Thus, this tracking technique recorded student comments, their navigational paths in the program and provided a chronological record of selections; all of which are useful for design purposes and for monitoring user-computer interactions.

Table 1
Selected Data: Imposition's Tracking System

Screen	User	# of Times on Screen	Sec. on Screen	Notes Comments
TitleScreen	1	1	20	
Purpose	1	1	34	
Help	1	1	26	
Purpose	1	2	9	
Menu	1	1	14	I do not like the
DummyFold1	1	1	12	background
DummyFold2	1	1	2	pattern of the
Animate	1	1	20	menu area.
ExeStep3	1	1	22	
Step 4	1	1	8	
Exercise Y/N	1	1	1	Are we to number
Binding	1	1	13	pages? I missed
DummySum	1	1	18	the directions.
Menu	1	2	4	
Menu	1	4	62	

Options:	Glossary	Notes	Comments
	0	0	2

When using Imposition, students typed comments which gave them the opportunity to express their likes and dislikes. These on-line capabilities permitted students to take notes, make comments and pose questions to the designer or express concerns while working. While some comments are more valuable than others, student narratives during formative evaluation helped effect many improvements in the software. In this case, they gave evidence of student frustration, confusion, likes, dislikes, ideas, etc. Collecting student narratives turned out to be an effective way to obtain data about the design of the module. The comments feature was useful for getting direct feedback from users in narrative form while they worked.

Video Recording (Video Split-Screen)

In assessing multimedia, potential users are often asked to explore a prototype and to write down any reactions they have to it. Paper-based written narratives, while useful, may provide limited information for several reasons. First, collecting written assessments (a laborious task) places the burden of evaluation on the subject. Thus, the richness of feedback may be compromised based on whether or not the subject is motivated to actively participate in evaluation. Moreover, depending upon what is being studied, written assessments may jeopardize the type of mental processing that can be observed. For example, a written assessment may not fully reflect a subject's instantaneous processing of stimuli or the way in which one explores and links conceptual events presented by the learning system. Second, subjects often fail to note all their reactions and comments or make critical observations. Third, the value of a written narrative is largely contingent upon an individual's ability to clearly communicate. Ambiguities can be left open for the researcher's interpretation. Fourth, immediate responses or physical and verbal expressions are frequently lost with written assessment.

With current technologies, multifaceted approaches to evaluation which include audio and video records have become possible. These provide researchers instruments by which to improve evaluations methodologies, as well as to study human cognition. The computer's ability to record students' actions combined with visual and auditory records such as video depictions of students working and thinking aloud can give insight into learners' cognitive processes (Ericsson & Simon, 1982). "The integration of computer and video records allows for powerful analyses of qualitative data, and the sharing of these analyses among researchers." (Kosma, 1991, p.206).

In determining methods by which to evaluate interactive multimedia learning systems and to study students' mental processes in such environments, Gibbs and Shapiro (1993) developed a video split-screen data collection method. This method simultaneously video records subjects' actions and decisions in a program, their verbal commentary and elaborations, observations, and nonverbal and attitudinal reactions. Subject and computer screen are each simultaneously video recorded. One camera records the subject and the other records the computer screen. The two images are integrated through a video effects generator to create a split-screen effect. This technique provides a permanent visual and auditory account of evaluation for subsequent analysis.

The split-screen method was effective for assessing a multimedia prototype. The video recordings allowed the researchers to visually and aurally reconstruct the actions taken by subjects. This is beneficial for several reasons. First, it enables one to simultaneously monitor, among other things, how a program is used, the types of options selected and the order in which they are selected. It also shows where learners appear confused, what they like and dislike, and their observations about the program. Second, the technique allows for learners' direct feedback about their interactions with the program. Third, it enabled program developers to readily identify and correct program execution errors by reconstructing the actions taken by learners. Fourth, it permits the review and validation of observations by several researchers.

Video-Split Screen and Cognitive Processing

Our understanding of learners' thought processes in various types of computer-based learning environments is limited (Nakhleh & Krajcik, 1991). The video split-screen method may aid researchers in studying mental processes. Cognitive information processing and cognitive views of knowledge construction provide a theoretical framework to support this claim. In explicating how this might be so, it is useful to review aspects of cognitive information processing and knowledge construction.

Cognitive Information Processing

An information processing perspective of cognition is categorized by transformation of information from stimuli in the environment to a response by the learner. The process is initiated as receptors receive information in the form of stimulation from the outside world. These stimuli get transmitted as information to the central nervous system. For a brief period, the information is recorded in the sensory registers and a minute fraction of it is sent on to Short-Term Memory (STM) and all remaining information is lost. STM has a limited information storage capacity and information, unless rehearsed, passes out of it in a very short time. STM is referred to as working memory which signifies the process of information encoding as new information becomes integrated with learned knowledge. Information in STM is accessible. Encoded information gets stored in Long-Term Memory (LTM) for extended periods of time. For information in long-term memory to be used, it enters working memory by a process known as retrieval. As new information enters working memory, the spread of activation prompts the retrieval of existing related knowledge. Through these processes, existing knowledge becomes integrated with incoming information to facilitate new learning (Gagne, 1985; Woolfolk, 1990).

The information processing model indicates that information in STM is accessible for verbalization and thus learners' verbal reports may indicate the contents of STM and/or the information to which learners attend (Nakhleh & Krajcik, 1991). Correspondingly, verbalizations made during an activity have been shown to reflect the thoughts in STM (Ericsson and Simon, 1984). Given this, the video-split screen along with think aloud verbal protocol analysis (e.g., Ericsson & Simon, 1984) may prove useful in understanding learners thought processes. Structured observations and protocol analysis as used by Nakhleh and Krajcik (1991), is a process whereby learners, when performing an activity are prompted to talk and think aloud. Their performance is video recorded and verbal commentaries subsequently transcribed. Each spoken phrase is coded with categories for protocol analysis suggested by Ericsson and Simon (1984).

A pilot study was conducted using the video split-screen and protocol analysis. Five subjects, all of whom were computer novices, were video recorded using identical content from a multimedia system. Evaluation sessions varied from 45 minutes to 1 hours and 15 minutes. Individually, subjects received a brief introduction to the evaluation session and the program. They were told to think aloud as they explored the system, to comment on it, to make observations, and to ask questions when necessary. Subjects used the system individually and could terminate it at any time. Selected verbal commentaries were reviewed and categorized and category frequencies tallied. The categorization scheme used (see Table 2) was a modification of that which Nakhleh and Krajcik (1991) employed in similar inquiries. Table 3 shows the way in which verbal commentaries were recorded. This table reflects a sample of categorizations for each of the 5 subjects' commentaries for approximately a 2.5 minute period at the beginning of their session.

While the selected reviews were from a very small sample and the categorizations were not checked for evaluator bias, the cursory review provides some indications as to the utility of such an analysis. For example, although viewing the same system and content, subjects were found to verbalize differently and for varying amounts. Accordingly, the multimedia systems was a prototype and required design modifications and thus many commentaries focused on procedural and interface issues related to the program (e.g., "When can I proceed?" "There were no other directions so I hit OK!"). Conversely, there were far fewer analytical statements (e.g., "The numbers are being highlighted in some random order and I don't know if there is a reason for that. I think the mouse cursor follows -- I see how it follows. It follows in a horizontal or vertical pattern.").

Table 2
Coding Categories

1. Procedural statements referring to (P):
    a. Reading or questioning directions
    b. Performing an action
    c. Stating a goal
    d. Deciding what to do next or admitting not knowing what to do next
2. Analytical statements referring to (A):
    a. Observing, interpreting, or explaining event or text.
    b. Understanding or not understanding observations or text.
    c. Hypothesizing about concepts.
    d. Recalling pertinent subject matter knowledge.
3. Emotional statements referring to (E): Puzzlement, frustrations, or satisfaction.
4. Statements of inadequate understanding of (S1): subject matter concepts
5. Statements of adequate understanding of (S2): subject matter concepts

(Modified from Nakhleh & Krajcik, 1991, p.11)

If it is assumed that verbalization reflect thoughts in short-term-memory, it appears that much of the processing, in this case, focused not on instructional content but rather program design and content access issues. Thus, if the purpose of the technology-based learning environment is to engage learners in higher order thinking and problem solving, then by using these techniques researchers may, with greater likelihood, determine the extent to which the system meets its purpose. This technique provides a visual representation and an auditory categorization scheme to profile learners thought processes and focus of attention. It may be possible to visualize a learners' navigation of the problem space, to monitor their thinking, the system's affect on thinking and the kinds of processes.

Table 3
Categorizations for Subject Commentaries (2.5 minute period)

(P): Procedural statements (A): Analytical statements
(E): Emotional statements (S1): Statements of inadequate understanding
(S2): Statements of adequate understanding

Ss 1	Ss2	Ss3	Ss4	Ss5
P	P	E	E	E
P	P	P	S2	A
P	P	S1	A	A
P	S1	S1	P	A
P	S1	S1	A	A
P	E	P	S1	E
P	E	P	S1	P
S1	S1	P	S1	P
A	S1	P	S1	P
P	P	S1	E	E
P	S1	P	P	S2
P	P	S1	P	
P	P	S2	P	
A	P	P	P	
A	P	S2	P	
A	A	P	P	
P	S1	P	P	
S1	S1	P	P	

used to derive solutions. As Nakhleh and Krajcik (1991) note, "...videotaped records capture the correspondence between students' observed actions and their verbal commentary about their prediction, explanations, observation, and procedural decision" (p.3).

Knowledge Construction

Based on cognitive principles of psychology, learning is no longer viewed as a passive process but an active one in which learners generate their own knowledge. Learning occurs as a results of students building their own cognitive structures (Wittrock, 1986, 1974) based on their background, experiences and attitudes. During learning, learners interact with their environment in constructing their own understanding of a subject (Jonassen, 1995). It is the intent of many designers to create environments in which learners are meaningfully engaged in ways that facilitate knowledge construction. If it is assumed that learners do create their own knowledge, then the instructional events in which they engage and their cognitive processes during these events become important (Nakhleh & Krajcik, 1991). Thus, analysis of learners cognitive processes during an instructional event (e.g., analyzing new information based on prior knowledge, modifying existing knowledge constructs) should be useful in more thoroughly understanding how they construct meaning, as well as how to design environments to facilitate these processes. With video split-screen and protocol analysis, it is possible to identify the type of thought processes in which learners engage. For example, by analyzing verbalizations, one may determine if and how a learner is accessing existing knowledge to integrate new information with it. More importantly perhaps, is the ability to assess how the system interferes or alters processing and/or how it is used based on individual thought processes. In this respect, the video split-screen and verbal protocol analysis, a few of the many possible approaches, can provide researchers insight in the design and evaluation of learning systems, and prove useful in the study of learners' thought processes.

Summary

This paper has discussed several approaches for assessing multimedia and hypermedia learning systems. To some degree, non-sequential hypermedia systems have increased the complexity by which to monitor the human-computer interface. Recent assessment approaches have become multifaceted and holistic in an attempt to more fully understand this interface. Data collection instruments such as computerized tracking combined with video recording techniques have the potential to provide more direct comprehensive evaluations. Accordingly, these approaches also provide researchers instruments by which to study cognitive processes. The cognitive information processing model and learning as a constructive idiosyncratic process provide a theoretical basis for directing the utilization of instruments such as those discussed in this paper to study learners' mental processes.

References:

Allessi, S.M. & Trollip, S.R. (1985). Computer–based instruction: methods and development. Prentice–Hall Inc., Englewood Cliffs, New Jersey.

Card, S., Moran, T., & Newell, A. (1983). The psychology of human-computer interaction. Hilllsdale, NJ: Lawarence Erlbaum Associates.

Ericsson, K., & Simon, H. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT press.

Gagné, E. (1985). The cognitive psychology of school learning. Boston: Little, Brown and Company.

Gay, G., & Mazur, J. (1993). The Utility of Computer Tracking Tools for User-Centered Design. Educational Technology, 33(4) 45-59.

Gibbs, W.J. & Armel, D.J. (1994). A Computer-Based Self-Study Module: An Analysis of Student Reactions. Proceedings of the 27th Annual Association of Small Computer Users in Education Conference. Myrtle Beach, South Carolina, June.

Gibbs, W.J. & Shapiro, A.F. (1994). Video split-screen technology: A data collection instrument. Journal of Computing in Higher Education, 5(2), 113-121.

Hannafin, M. J., & Peck, K. L. (1988). The design, development, and evaluation of instructional software. New York: Macmillan Publishing.

Jonassen, D.H. (1995). Computer as cognitive tools: Learning with technology, not from technology. Journal of Computing in Higher Education, 6(2), 40-73.

Kozma, R.B. (1991). Learning with media. Review of Educational Research. 61(2), 179-211.

Kumar, D.D., Helgeson, S.L., & White, A.L. (1994). Computer technology-cognitive psychology interface and science performance assessment. Education Technology Research and Development, 42(4), 6-16.

Marchionini, G. (1990). Evaluating hypermedia systems. In Proceedings of the CH '90 Conference on Computer and Human Interaction. (pp. 387-390). April 1990, Seattle. New York: ACM.

Nakhleh, M. B., & Krajcik, J. S. (1991). The use of videotape to analyze the correspondence between the verbal commentary of students and their actions when using different levels of instrumentation during laboratory activities. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Lake Geneva, WI.

Park, I., & Hannafin, M. J. (1993). Empirically-based guidelines for the design of interactive multimedia. Education Technology Research and Development, 41(3), 63-85.

Wadlow, M.G. (1990). The andrew system: The roles of human interface guidelines in the design of multimedia applications. Current Psychology: Research & Reviews, 9(2), 181-191.

Winograd, T., & Flores, F. (1986). Understanding computer and cognition: A new foundation for design. Reading, MA: Addison-Wesley.

Woolflok, A. E. (1990). Educational Psychology. Englewood Cliffs, New Jersey: Prentice Hall.

Wittrock, M.C. (1986). Student thought processes. In M.C. Wittrock (Ed.), Handbook of Research on Teaching (3rd ed., pp. 297-314). New York: Macmillan.

Wittrock, M.C. (1974). Learning as a generative process. Educational Psychologist. 11, 87-95.